Amendments to the Claims

1. (currently amended) A method for detecting highlights from videos, comprising:

extracting audio features from the video;

classifying the audio features as labels;

extracting visual features from the video;

classifying the visual features as labels; and

fusing, probabilistically, the audio labels and visual labels into a single discrete-observation coupled hidden Markov model to detect highlights in the video.

2. (currently amended) The method of claim 1, in which the video is compressed, and the single discrete-observation coupled hidden Markov model includes the audio features, the visual features, audio states of the audio features and visual states of the visual features.

3. (original) The method of claim 1, in which silent features are classified according to audio energy and zero cross rate.

4. (original) The method of claim 1, in which the audio features are Mel-scale frequency cepstrum coefficients.

5. (original) The method of claim 1, in which the audio features are MPEG-7 descriptors.

6. (original) The method of claim 1, in which the audio features are classified using Gaussian mixture models.

7. (original) The method of claim 1, in which the audio labels are selected from the group consisting of applause, cheering, ball hit, music, male speech, female speech, and speech with music.

8. (original) The method of claim 1, in which the visual features are based on motion activity descriptors.

9. (original) The method of claim 1, in which the visual features include dominant color and motion vectors.

10. (original) The method of claim 1, in which a variance of the motion activity is quantized to obtain the visual labels.

11. (original) The method of claim 1, in which the motion activity is averaged to obtain the visual labels.

12. (original) The method of claim 1, in which the visual labels are selected from the group consisting of close shot, replay, and zoom.

13. (canceled)

14. (currently amended) The method of ~~claim 13~~ claim 1, in which the discrete-observation coupled hidden Markov model includes audio hidden Markov models and visual hidden Markov models.

15. (original) The method of claim 14, in which the discrete-observation coupled hidden Markov model is generated from a Cartesian product of states of the audio hidden Markov models and the visual hidden Markov models, and a Cartesian product of observations of the audio hidden Markov models and the visual hidden Markov models.

16. (currently amended) The method of ~~claim 13~~ claim 1, further comprising:

training the discrete-observation coupled hidden Markov model with hand labeled videos.

17. (original) The method of claim 1, in which the video is a sport video.

18. (original) The method of claim 1, further comprising:

determining likelihoods for the highlights; and

thresholding the highlights.

19. (currently amended) The method of claim 2, in which transitions between the audio states and the visual states of the single discrete-observation coupled hidden Markov model are represented by transition matrices ~~portion of the video is compressed~~.

20. (currently amended) The method of claim 1, in which the probabilistic fusion is according to a function fusion function $F(f_A, f_B)$, where f_A are the audio features and f_B are the visual features ~~visual portion of the video is compressed~~.

21. (new) The method of claim 19, in which the transition matrices have a form:

$$a^1_{(i,j),k} = Pr(S^1_{t+1} = k | S^1_t = i, S^2_t = j)$$
$$1 \leq i, k \leq M; 1 \leq j \leq N$$

where S^1 represents the audio states and S^2 the visual states, and Pr is a probability of transitioning to a next state k at the next time instant t given two current hidden states i and j, respectively, and M is a number of audio states and N a number of visual states.